Testing for Asymmetric Employer Learning and Statistical Discrimination

Suqin Ge, Andrea Moro, and Beibei Zhu*

June 4, 2020

Abstract

We test if firms statistically discriminate workers based on race when employer learning is asymmetric. Using data from the NLSY79, we find evidence of asymmetric employer learning. In addition, employers statistically discriminate against non-college educated black workers at time of hiring. We also find that employers directly observe most of the productivity of college graduates at hiring and learn very little over time about these workers.

Keywords: statistical discrimination, employer learning, asymmetric learning
JEL code: J71, D82, J31

1 Introduction

In an influential paper, Altonji and Pierret (2001) (AP hereafter) adopted the landmark model of employer learning by Farber and Gibbons (1996) to test whether employers statistically discriminate workers by race. In this line of research, learning about workers' productivity occurs over time, after observing signals of workers' productivity. Time is included in the empirical specification as workers' experience. The implicit assumption is that outside employers, when attracting workers, have the same information about workers' skills as the workers' current employer. In this paper, we drop this assumption and test for statistical discrimination based on race

*We would like to thank Luca Flabbi, Limor Golan, John Kennan, Kevin Lang, Ben Rosa, seminar and conference participants from various institutions for their valuable comments and suggestions. This study benefited from funding from the U.S. Department of Labor. The views expressed are those of the authors and should not be attributed to the Federal Government or the Department of Labor. Contact information: Ge, Department of Economics, Virginia Tech, E-mail: `ges@vt.edu;` Moro, Department of Economics, Vanderbilt University, E-mail: `andrea@andreamoro.net;` Zhu, Slack, E-mail: `zhubeibei0319@gmail.com`.

while assuming that learning can occur asymmetrically, that is, outside employers may have less information on productivity than current employers.

In statistical discrimination models, productivity and qualifications of labor force participants are difficult to observe directly. Therefore imperfectly informed employers use demographic characteristics, such as race or gender, as proxies for unobserved worker characteristics.[1] A number of studies provide empirical evidence showing that employers learn over time (in addition to AP, see e.g., Lange (2007), Arcidiacono et al. (2010), and Mansour (2012)). A direct implication of employer learning is that firms become less inclined to statistically discriminate based on observed group characteristics as they accumulate over time additional information about individual workers' productivity. Hence, employers rely less and less on group characteristics as proxies for productivity over time, and wages become more correlated with measures of productivity available to the investigator. However, existing studies on employer learning and statistical discrimination are carried out under the assumption that employer learning is symmetric, that is, incumbent and outside firms have the same information about workers' productivity.[2]

Some theoretical articles have studied the hypothesis of "asymmetric employer learning," that is, current employers are at an informational advantage about workers' productivity than outside employers. In this literature, asymmetric information between outside and current employers generates market power for the current employer, breaking the link between expected productivity and wages that is used for identification in AP and related literature.[3] There is no comprehensive theory of the precise nature of the relationship between wages and asymmetric learning and

[1] The two main branches of statistical discrimination theories are screening discrimination and rational stereotyping. The former, originated from Phelps (1972), attributes discriminatory outcomes to unexplained exogenous differences between groups, combined with employers' imperfect information about workers' productivity. This literature (see also Aigner and Cain (1977)) is largely agnostic on where the initial group differences originate. They may result from either differences in employer perceptions or other factors, such as differences in the quality of education or human capital acquisition. The other branch of this theory, originated from Arrow (1973) and modeled most comprehensively by Coate and Loury (1993), assumes that employer's negative beliefs about the quality of minority workers are self-fulfilling and thus average group differences are endogenously derived in equilibrium. Fang and Moro (2011) provide a detailed survey on the theoretical literature on statistical discrimination, and Lang and Lehmann (2012) offer an extensive survey on theory and empirics of racial discrimination.

[2] Most of the studies focus on males using U.S. data. A notable exception is Lesner (2018), who finds evidence of statistical discrimination against women using a Danish sample.

[3] The details of the relationship between wages and expected productivity depends on specific modeling assumptions. In Pinkston (2009)'s model, for example, a bidding process between outside and current employers is assumed which completely reveals the current employer's information. In other models, stemming from Waldman (1984)'s seminal contribution, outside employers observe the worker's job rank. Because they can make inference from promotions, this generates inefficiencies in worker's assignment to job ranks. Other examples of theoretical models of asymmetric employer learning include Greenwald (1986), Bernhardt (1995), Golan (2005), and Golan (2009).

its implications in presence of statistical discrimination. While the characterization of these relationships is an interesting avenue for future theoretical research, in this paper we focus on testing specific empirical hypotheses which hold regardless of the modeling details.

First, under symmetric learning, the learning process occurs over a worker's general experience regardless of job turnovers. By contrast, if outside firms have no information on a worker's productivity, then the learning will change once the worker moves from one job to another. Therefore, the learning process takes place over job tenure rather than general experience. As a consequence, the correlation of wages with measures of skills observed by the econometrician should increase more with tenure than with experience, and the opposite occurs when learning is symmetric. Secondly, we also expect that when employers statistically discriminate against minorities, the initial wages of minority workers (conditional on skills) are lower than other workers' wages whenever workers finds a new job. As employers learn about productivity over time, the effects of race decrease conditional on measures of skills available to econometricians.

We test these implications using the National Longitudinal Survey of Youth 1979 (NLSY79), the same data used in AP, but including more recent waves. We follow the literature in using the standardized value of the Armed Forces Qualification Test (AFQT), a battery of aptitude tests, as the measure of skills observed by the econometrician. Using the sample of non-college educated workers, we find evidence that employers learn asymmetrically about workers' skills and that they statistically discriminate against black workers. Wages become more correlated with skills as time passes, and this correlation increases more when tenure is used as a measure of time, as opposed to experience, consistent with asymmetric learning. Moreover, black workers without a college education suffer a wage penalty initially, but wages become more correlated with skills over time.

Results are different for college educated workers. For this class of workers we find neither evidence that learning is asymmetric nor evidence of statistical discrimination against black workers. We conjecture that key aspects of worker productivity are directly observed by employers upon initial labor market entry; as such, little learning takes place subsequently, consistent with the main findings reported in Arcidiacono et al. (2010).

Our paper relates to other empirical papers that test asymmetric learning, but do not focus on its implications regarding statistical discrimination. This empirical literature offers no conclusive evidence on the nature of employer learning. Schönberg (2007) studies a two-period model of asymmetric learning and derives implications for job transitions and wage dynamics. Using a sample of white males only, she finds that employer learning is mostly symmetric. Pinkston (2009) also tests the implications of asymmetric learning and finds that asymmetric employer learning plays a role that is at least as important as symmetric learning. Kahn (2013) investigates asymmetric learning by using an original approach that looks at the implications on the variance of wage changes for workers that change jobs and workers that stay with their current jobs. She finds support for asymmetric information between incumbent and external employers but does not focus on racial differences. Other studies (Gibbons and Katz (1991), Bauer and Haisken-DeNew (2001)) find empirical evidence in favor of asymmetric employer learning, that is, current firms have access to more information about workers' productivity than outside firms. Fan and DeVaro (2020) find that job hopping is associated with lower wages for college graduates, suggesting asymmetric learning. Finally, Bates (2020) finds evidence of asymmetric learning of teachers' effectiveness using value added measures of student's achievements. Our contribution, relative to this literature, is to focus on the implications of asymmetric learning on racial differences generated by statistical discrimination and to test them by educational level.

The remainder of the paper is structured as follows. Section 2 outlines our theoretical framework. Section 3 describes our empirical specifications. Section 4 provides an overview of the data and compares the results in AP with the results from our sample by using the same specification. Sections 5 and 6 present the main results. Section 7 concludes and suggests directions for further research.

2 Theoretical Effects of Employer Learning and Statistical Discrimination

Before turning to estimation, it is useful to outline the ways in which employer learning and statistical discrimination can affect the evolution of wages over time.[4] We consider an environment where firms cannot directly observe labor force participants' true productivity and qualifications but they may learn over time. Employers initially observe the characteristics (such as gender, race and education) of a new worker and some information (an initial signal) on the worker's productivity, such as his communication skill and ambition, which employers can (at least partially) observe through job interviews. After hiring, new information (additional signals) about the worker's job performance becomes available to his employer, and the em-

[4] In the External Appendix, after discussing the theoretical challenges, we use a simple signal extraction model to illustrate the implications presented in this section.

ployer learns about the worker's productivity over time.

In each period, the employer computes the worker's expected productivity given the information observed. When the worker enters the labor market for the first time, the employer has incomplete information about the worker's productivity and uses demographic characteristics, such as gender, race or education, as proxies for unobserved worker productivity. Expected productivity is a weighted average of the population (based on observed characteristics) average productivity and the initial signal. As the worker increases his tenure with a firm, the incumbent employer exploits new information from multiple signals, providing more precise information about true productivity. Hence, the employer relies less on the population mean and puts more weight on productivity signals in predicting the worker's productivity over time.

A crucial assumption made in the employer learning and statistical discrimination literature is that the researcher has access to a correlate of workers' productivity that is unobserved by employers. Altonji and Pierret (2001) and the line of research following their study assume that AFQT score is such a proxy correlate of productivity.[5] Although each worker's individual AFQT is unknown to employers, employers observe the average AFQT of each group(for example, by race or education). Under the assumption that wages are determined by expected productivity, an empirical implication of the employer learning model is that wage becomes more correlated with AFQT over job tenure.

Next we consider the evolution of the correlation between wage and AFQT over time when workers hold multiple jobs over their lifetime. After a worker is hired by a new employer, the new employer collects information from the worker's resume and other signals. If the information about workers' productivity available to the new employer at the time of hiring is the same as the information available to the current employer, then employer learning is symmetric (or public). In contrast, if the information available to the new employer is worse than the information available to the current employer, then employer learning is asymmetric. We believe the latter to be more realistic because worker's resume and job interviews cannot substitute from day-to-day interactions over the workers' tenure. If the new employer does not infer any information from prior job history, we label the learning process as *purely asymmetric*. The different natures of employer learning suggest different predictions about how wages evolve with job experience versus job tenure. Symmetric learning implies a continuous learning process over a worker's general labor market experience

[5] Lange (2007) offers a detailed discussion on a number of reasons why employers do not observe the AFQT score, including the high turnover of employees, which limits the economic returns to administering such tests, and managers' concerns about legal repercussions.

regardless of job turnovers. If learning is asymmetric, then employer learning will be interrupted once the worker moves from one job to another, and the learning process takes place more over job tenure than over general experience. Therefore, the correlation of wage with AFQT would increase more with tenure than with experience.

Workers belong to an identifiable group differing in race, gender, education, etc. Employers observe group membership and may use the group membership as a cheap source of information about productivity because of the perceived correlation between group membership and test scores such as AFQT. A statistical discriminating employer may use group membership, such as race, along with other information to predict workers' productivity at the time of hiring. Over time, the employer observes more signals about the productivity of the worker and thus relies less on the group membership information available initially. In a wage regression including a minority group dummy, if the minority group is statistically discriminated against, then the coefficient on such dummy is negative, but its interaction with time is positive so that the negative effect declines over time. If learning is symmetric, the interaction between group dummy and experience should be positive as learning takes place over experience; whereas if learning is asymmetric, the interaction between group dummy and tenure should be positive because learning takes place primarily over tenure.

3 Empirical Specification

We propose empirical specifications motivated by the theoretical framework presented above. If employer learning is symmetric, then the learning process occurs over general work experience regardless of job turnovers. By contrast, purely asymmetric learning implies that only current employers learn about workers' productivity over time, so that learning only takes place over job tenure. To distinguish the two learning hypotheses, we use actual work experience X and job tenure T as two separate time measures. We estimate the following corresponding wage equations:[6]

$$\begin{aligned}\ln w_i = \beta_0^X + \beta_S^X S_i + \beta_{S,X}^X (S_i \times X_i) + \beta_{AFQT}^X AFQT_i + \beta_{AFQT,X}^X (AFQT_i \times X_i) \\ + \beta_{Black}^X Black_i + \beta_{Black,X}^X (Black_i \times X_i) + \beta_\Omega^X \Omega_i + H(X_i) + \epsilon_i^X, \qquad (1)\end{aligned}$$

[6]In general, learning is nonlinear in time, which implies that the effects of AFQT score and race should also vary nonlinearly with time. For simplicity, however, we follow the literature and assume the relationships between log wage, AFQT score, and race to be linear in time.

$$\ln w_i = \beta_0^T + \beta_S^T S_i + \beta_{S,T}^T (S_i \times T_i) + \beta_{AFQT}^T AFQT_i + \beta_{AFQT,T}^T (AFQT_i \times T_i)$$
$$+ \beta_{Black}^T Black_i + \beta_{Black,T}^T (Black_i \times T_i) + \beta_\Omega^T \Omega_i + H(T_i) + \epsilon_i^T. \quad (2)$$

where w_i denotes the hourly wage of individual i, S_i indicates the years of schooling, $AFQT_i$ denotes individual AFQT score, $Black_i$ is a dummy variable on race, and Ω_i is a vector of demographic variables and other controls. In all of our specifications, we control for urban residence, dummies for region of residence, and year fixed effects. The variables X_i and T_i measure time, and $H(\cdot)$ is a polynomial in time. Time is measured in months in our sample, and we divide the interaction of any variable with time measure by 120; thus the coefficients on interaction terms measure the change in wage during a 10-year period. In the empirical analysis below, we follow the literature and assume the effects of AFQT and *Black* on log wages to vary linearly with time to simplify the interpretation of these coefficients.

Our discussions from last section suggest that the coefficient $\beta_{AFQT,T}^T$ should be positive, and $\beta_{AFQT,T}^T$ should be significantly larger than $\beta_{AFQT,X}^X$. Under statistical discrimination, $\beta_{Black}^T < 0$ and $\beta_{Black,T}^T > 0$ when learning is asymmetric.

4 Data

The empirical analysis is based on the 2008 release of NLSY79, a nationally representative sample of 12,686 young men and women who were 14–22 years old when they were first surveyed in 1979. These individuals were interviewed annually until 1994 and on a biennial basis thereafter. The dataset contains detailed information on family background, academic performance, and labor market outcomes of a cohort of young workers; moreover, the weekly work history data provide information to construct accurate measures of actual work experience and job tenure.

The empirical analysis is restricted to black and white male workers who have completed at least eight years of education, thus we use the same restriction as in AP. We only analyze labor market observations after a person makes school-to-work transition. An individual is considered to have entered the labor market when he leaves school for the first time. Following the criteria used by Arcidiacono et al. (2010), military jobs, self-employed jobs, jobs at home, and jobs without pay are excluded from the construction of experience and from the analysis because we want to focus our analysis on civilian employees.

We derive individual monthly employment status by using work histories, which contain each respondent's week-by-week labor force status since January 1978. An individual is considered employed in a given month and accumulates one month

Table 1: Summary Statistics by Race

	Whites			Blacks		
	All	<College	≥College	All	<College	≥College
AFQT	0.501	0.201	1.345	-0.571	-0.726	0.482
	(0.957)	(0.885)	(0.568)	(0.796)	(0.649)	(0.897)
Education (yrs)	13.35	12.14	16.74	12.69	12.12	16.60
	(2.39)	(1.31)	(1.19)	(2.00)	(1.35)	(1.06)
Hourly wage	12.91	11.10	17.99	10.15	9.23	16.41
	(8.14)	(5.89)	(10.97)	(6.14)	(4.98)	(8.96)
Experience:						
Potential	131.41	135.07	122.12	145.37	147.71	129.46
	(84.51)	(86.20)	(78.66)	(85.76)	(86.50)	(78.73)
Actual	110.54	111.97	106.52	111.15	111.28	110.25
	(76.05)	(77.50)	(71.68)	(73.17)	(73.66)	(69.76)
Job tenure	46.94	45.98	49.63	40.90	40.17	45.92
	(48.42)	(48.29)	(48.70)	(43.30)	(42.76)	(46.50)
Individuals	2,592	1,906	686	1,133	987	146
Observations	224,304	165,480	58,824	93,684	81,660	12,024

Notes: Standard deviations are in parentheses. Education is measured in years, real hourly wages in 1990 dollars, and experience in months. Potential experience is months since left school.

of work experience or tenure if he works at least 10 hours per week for at least three weeks, or during the last two weeks of the month. Otherwise, an individual is classified as nonemployed. The work history information is employer-based, thus a "job" should be understood as an uninterrupted employment spell with an employer. We link all jobs across survey years and build a complete employment history for each respondent in the sample. Multiple jobs held contemporaneously are treated as a new job, with an associated wage equal to the average wage weighted by hours on each job, and working hours equal to the sum of working hours on the different jobs. Tenure on a job is completed when an individual makes a job-to-job transition or when she is back in non-employment. Job tenure is the number of months between the start of a job and either the date the job ends or the interview date. Actual work experience is the sum of tenure for all jobs.[7] Potential work experience is defined as months since the respondent first left school.

The wage measure that we use is the hourly rate of pay on each job, provided in the work history file. Nominal wages are deflated to real hourly wages in 1990 dollars by using the monthly CPI released by the BLS. We exclude observations with

[7]In AP, actual experience is defined as the weeks worked divided by 50. Our measure is very close to theirs and more compatible with our tenure measure.

real wages less than $1 or more than $100 per hour. We use the AFQT as our proxy correlate of productivity. To eliminate age effects, we standardize the AFQT score to have a mean zero and standard deviation one for each three-month age cohort. We use data from the main cross-sectional sample of the NLSY79 and the supplementary sample, which oversamples blacks and disadvantaged whites.[8] The total remaining sample consists of 2,592 whites and 1,133 blacks with 317,988 monthly observations. We also consider in the analysis two education samples: white or black men who have completed at least 16 years of education (college graduates sample) or less than 16 years of education (non-college graduates sample).[9]

Table 1 presents the summary statistics for the main variables in our sample by race and education level. The average AFQT score of black workers is about one standard deviation lower than that of white workers, possibly as a result of pre-market discrimination or racial bias in testing. This test score gap persists even if we control for education. Black workers generally earn lower wages and accumulate less job tenure than white workers. Potential employers have strong incentives to statistically discriminate on the basis of race if AFQT is a good measure of skill. In the next section, we carry out the empirical analysis to examine this issue in detail.

In Table 2, we compare the results by using different samples. We report for convenienence in column (1) the results from AP's Table 1, panel 1, column 4. Despite the differences, the main qualitative results from AP are confirmed. The coefficient on education is positive and significant initially and falls over time. The coefficients on AFQT and AFQT–experience interaction imply that the impact of AFQT score on log wages rises as workers accumulate experience. That is, employers learn about workers' productivity over time, so the weight they put on the hard-to-observe correlate of productivity, AFQT, increases. The coefficient on *Black* is small and not significant at the time of initial hire, but it becomes significantly negative over time. Given that the racial wage gap is initially not statistically different from zero, AP conclude the lack of statistical discrimination on the basis of race.

The specification in column (2) uses data from the same time period (interview years 1979–1992), but with some differences in sample construction that we adopt in our analysis. First, we use monthly data instead of annual data. Second, we measure potential experience as time since first left school instead of age minus years of schooling minus 6. Column (3) reports analogous results by using our full

[8]All statistics in this study are unweighted. Using sampling weights does not change the qualitative results.

[9]Considering high school dropouts and workers with some college education but without a college degree behave similarly to high school graduates, we bundle them into a sample of workers with no college degree.

Table 2: Sample Comparisons

Time	1979–1992	1979–1992	1979–2008	1979-2008
	(1)	(2)	(3)	(4)
Education	0.079***	0.088***	0.071***	0.080***
	(0.015)	(0.007)	(0.006)	(0.006)
Education × experience/120	-0.019	-0.035***	-0.002	-0.018*
	(0.013)	(0.009)	(0.004)	(0.007)
AFQT	0.022	0.035*	0.057***	0.036**
	(0.042)	(0.014)	(0.012)	(0.013)
AFQT × experience/120	0.052	0.069***	0.037***	0.071***
	(0.034)	(0.018)	(0.009)	(0.014)
Black	-0.057	-0.030	-0.037	-0.039
	(0.072)	(0.026)	(0.022)	(0.025)
Black × experience/120	-0.083	-0.084**	-0.053***	-0.053*
	(0.058)	(0.031)	(0.015)	(0.026)
R^2	0.287	0.273	0.346	0.322
No. of Observations	21,058	177,288	317,988	212,640

Notes: Column (1) reproduced from AP's Table 1, Panel 1, Column 4. In columns (2)-(4), the experience measure is months since left school for the first time. All specifications control for year effects, urban residence, region of residence, experience, and experience squared. The numbers in parentheses are White/Huber standard errors that account for multiple observations per person. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

sample, the 1979-2008 waves of NLSY79. We obtain qualitatively similar results, but AFQT and *Black* now have flatter profiles with experience and the returns to AFQT are greater initially. The difference in the time paths of AFQT and *Black* is likely driven by a non-linear employer learning process. To make our sample more comparable with the AP sample, in column (4), we restrict our sample to experience level less than 168 months, which are the maximum months of potential experience in the AP sample. This restriction restores the lower initial AFQT effect and its steeper profile over time.

5 Results

An important finding in employer learning literature is that the employer learning process may vary across different educational groups. Arcidiacono et al. (2010) find that a college degree helps workers directly reveal key aspects of their productivity; thus, employer learning is more important for high school graduates. They argue that if all education levels are pooled in wage regressions, then the estimates can be biased and the results may be misinterpreted. Based on their results, we split our sample into college graduates and non-college graduates; that is, a person who has completed at least 16 years of education is considered a college graduate and otherwise a non-college graduate.[10] We use these two samples from the 1979–2008 waves of NLSY79 to test the main predictions of our learning-based statistical discrimination model.

Our empirical analysis has two main focal points. First, to distinguish between symmetric and asymmetric employer learning, we examine the initial coefficients on AFQT and their interaction terms with time when experience and tenure are used as two separate time measures in log wage Equations (1) and (2). Second, we investigate how the racial wage gap varies over time to examine whether or not employers statistically discriminate against black workers. If employers hold racial prejudice, then our learning-based statistical discrimination model predicts a large initial racial wage gap because employers base payments on race and a narrowing racial gap over time as the employers accumulate additional information on true productivity.

In the first two columns of Table 3, we report estimates of the wage regressions using the non-college graduate sample. If employer learning is symmetric, then learning takes place over general work experience. The specification in column (1) estimates Equation (1) with actual work experience in months as the experience measure. We use actual work experience because it is a more accurate measure of workers' labor market experience than potential experience and the construction of actual experience and tenure are more consistent with each other. In the specification reported in column (2), we use tenure as time measure. The coefficients on AFQT and AFQT interacted with experience or tenure are all positive and significant, suggesting that productivity may be partially observed to employers at the time of initial hire and that employers learn about non-college workers' productivity over time as they acquire new information. The positive and significant AFQT-experience and AFQT-tenure interaction terms are consistent with the prediction of the employer learning model.

If employer learning is asymmetric, then learning takes place mostly on job tenure as outside firms have limited information regarding a worker's productivity. The coefficient on the AFQT-tenure interaction in specification (2) is greater than the estimated coefficient on AFQT-experience interaction in specification (1), with a P-value of 0.071. Overall, this evidence indicates that employer learn over time about

[10] Arcidiacono et al. (2010) restrict their sample to white or black men who have exactly a high school or a college degree with 12 or 16 years of education. If we restrict our college sample to those with 16 years of education and our high school sample to those with 12 years of education, then the empirical results are very similar to those we find below.

Table 3: Effects of AFQT and Race on Log Wages over Experience and Tenure

	Non-College Graduates		College Graduates	
	(1)	(2)	(3)	(4)
AFQT	0.051***	0.054***	0.123***	0.156***
	(0.012)	(0.010)	(0.026)	(0.024)
AFQT×experience/120	0.036***		0.038	
	(0.011)		(0.024)	
AFQT×tenure/120		0.065***		-0.036
		(0.018)		(0.041)
Black	-0.046*	-0.127***	0.138**	0.104*
	(0.021)	(0.019)	(0.047)	(0.046)
Black×experience/120	-0.042*		-0.091*	
	(0.020)		(0.038)	
Black×tenure/120		0.049		-0.155*
		(0.037)		(0.077)
R^2	0.258	0.253	0.268	0.262
No. of observations	247,140	247,140	70,848	70,848

Notes: the experience measure is actural work experience in months. All specifications control for years of education, year effects, urban residence, and region of residence. Specifications with experience also control for a quadratic term in actural experience and education interacted with experience, and specifications with tenure control for a quadratic term in tenure and education interacted with tenure. The numbers in parentheses are White/Huber standard errors that account for multiple observations per person. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

workers skills, learning re-starts at the beginning of each new job, and the speed of learning is faster over job tenure.

Turning to the analysis of racial differences, we find that at the time of initial entry into the labor market for non-college graduates and in contrast to the results from AP, black workers earn less than white workers with the same AFQT score in both specifications, supporting the hypothesis that employers have limited information about the productivity of new workers and statistically discriminate on the basis of race.[11] Although employer learning makes wages more correlated with skill over time, we do not find a strong evidence that the racial wage gap (conditional on measured skill) decreases over time: that is, the coefficients of race interacted with tenure or experience are either insignificant or negative.

[11] AP find little evidence on statistical discrimination on the basis of race and argue that statistical discrimination plays a relatively unimportant role in the racial wage gap. When we pool the education groups, we also find limited evidence on racial statistical discrimination. Mansour (2012) confirms AP's finding, but his empirical results imply that the pattern might differ across occupations.

In the last two columns of Table 3, we report the corresponding regression results for college graduates. The coefficients on AFQT are large and statistically significant but the coefficients on AFQT interacted with time are insignificant and relatively small in both specifications. These results are robust when we use actual experience and job tenure as alternative time measures as in columns (3) and (4). The time trend of the returns to AFQT shows substantial returns to AFQT for college graduate workers immediately after they take a new job. A one standard deviation increase in AFQT is associated with between 12.3%–15.6% increase in wages. Moreover, the returns to AFQT are not affected by experience or tenure. Following the interpretation by Arcidiacono et al. (2010), the estimated AFQT-time profiles suggest that employers have accurate information about the productivity of newly hired college graduate workers and learn very little additional information over time.

In contrast to non-college graduates, college-educated black workers earn higher wage than their white counterparts when they start a job, but this black wage premium declines over time.[12] Arcidiacono et al. (2010) argue that information contained on the resumes of college graduates, such as grades, majors, and college attended, help college-educated workers to directly reveal their productivity to their employers. Therefore, in the market for college graduate workers, employers have less incentives to statistically discriminate against black workers because they can assess workers' productivity more accurately at the time of initial hire. One plausible explanation for the black wage premium among college graduates is that black college workers are more motivated and productive than their white counterparts. If the AFQT and other tests, such as SAT, are racially biased, then blacks will have higher productivity than whites conditional on the test scores.[13] The diminishing black wage premium over time among high-skilled workers indicates that black workers may still suffer from racial prejudice in opportunities for promotion or on-the-job training over their careers despite the lack of statistical discrimination at hiring.

We conclude that employer learning mainly occurs with non-college graduate workers. Productivity is observed nearly perfectly for workers with a college degree at hiring; thus, limited scope is left for employer learning.

[12]The existence of a substantial black wage premium for college graduates is a robust feature of the U.S. labor market. Neal and Johnson (1996) find that the racial wage gap for males declines with the skill level, and a similar finding is also reported in Lang and Manove (2011).

[13]As argued by Arcidiacono (2005), affirmative action in the workplace may also account for the inital black wage premium. Black workers earn more because the number of blacks with a college of degree is small, yet employers value diversity in the workplace.

6 Additional results

6.1 Instrumental Variables (IV) Estimates

In specification (1), actual experience is determined by workers' employment decision, which may be correlated with individual productivity. This unobserved heterogeneity across individuals may produce inconsistent estimates of the effect of experience on wages and the speed of employer learning over experience. In addition, actual experience may be used by employers as a measure of quality (it is an indicator of the intensity of worker effort). Considering these potential endogeneity concerns, in column (1) of Table 4, we present the results from an instrumental variable (IV) specification, where actual experience is instrumented with potential experience for non-college graduates.

In Table 3, we also treat tenure as exogenous. However, tenure depends on quit and layoff decisions and may be correlated with characteristics of workers and job matches. These characteristics are likely to be related to worker productivity and how fast employers learn productivity. Therefore, we report in column (2) of Table 4 the results from an IV specification for non-college graduates. We use the variation of tenure over a given job match, following Altonji and Shakotko (1987), along with potential experience as instruments for job tenure. Specifically, our instruments are the deviations of the job tenure variables around their means for the sample observations on a given job match. This variable is by construction uncorrelated with individual and job specific unobserved components. Similarly, in columns (3) and (4) of Table 4, we report the corresponding IV estimates for college graduates where actual experience and job tenure are instrumented.

Overall, the IV estimates are very similar to the OLS estimates. The coefficients on AFQT and AFQT interacted with experience or tenure are all positive and significant, and the coefficient on the AFQT-tenure interaction is greater than the one on the AFQT-experience interaction for non-college graduates. This is consistent with asymmetric employer learning. The IV estimates also show that non-college black workers earn significantly less than white workers with the same AFQT, providing evidence on statistical discrimination on the basis of race. For college graduates, the estimated effects of AFQT and race on log wages by using IV specifications reported in Table 4 are also very close to those reported in Table 3. Therefore, our results are not driven by the potential endogeneity in work experience or job tenure.

Table 4: Effects of AFQT and Race on Log Wages: IV Estimates

	Non-College Graduates		College Graduates	
	(1)	(2)	(3)	(4)
AFQT	0.037**	0.053***	0.119***	0.133***
	(0.013)	(0.011)	(0.027)	(0.022)
AFQT×experience/120	0.052***		0.038	
	(0.011)		(0.025)	
AFQT×tenure/120		0.067***		0.025
		(0.015)		(0.032)
Black	-0.055*	-0.098***	0.163**	0.104*
	(0.023)	(0.019)	(0.051)	(0.043)
Black×experience/120	-0.042		-0.129**	
	(0.023)		(0.046)	
Black×tenure/120		-0.042		-0.159**
		(0.025)		(0.053)
R^2	0.253	0.251	0.257	0.261
No. of observations	247,140	247,140	70,848	70,848

Notes: the experience measure is actural work experience in months. All specifications control for years of education, year effects, urban residence, and region of residence. Specifications with experience also control for a quadratic term in actural experience and education interacted with experience, and specifications with tenure control for a quadratic term in tenure and education interacted with tenure. The numbers in parentheses are White/Huber standard errors that account for multiple observations per person. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

6.2 Non-Purely Asymmetric Learning

As a robustness check and to examine the possibility that employer learning is not purely asymmetric, we analyze how AFQT varies with experience and tenure when both are included in the regression model.[14] If learning is symmetric, then the coefficient on AFQT-tenure interaction should be zero because employer learning takes place over general experience. If learning is purely asymmetric, then outside firms are completely excluded from the learning process. We should only observed learning over tenure, thus AFQT-experience interaction should be zero. Otherwise if learning is not purely asymmetric, then some productivity information is revealed to outside firms but more information is available to current firms; thus, both AFQT-experience and AFQT-tenure interactions should have non-zero coefficients. We specify the following wage regression that includes both experience and tenure interaction terms

[14] Using a sample of white males from 1979–2001 waves of NLSY79, Schönberg (2007) examines whether employer learning is symmetric or non-purely asymmetric by analyzing how education and AFQT vary with experience and tenure when both are included in the wage regression.

Table 5: Testing for Non-purely Asymmetric Learning for Non-College Graduates

	(1)	(2)
AFQT	0.054***	0.046***
	(0.010)	(0.012)
AFQT×experience/120		0.018
		(0.013)
AFQT×tenure/120	0.065***	0.050*
	(0.018)	(0.022)
Black	-0.127***	-0.059**
	(0.019)	(0.021)
Black×experience/120		-0.065**
		(0.023)
Black×tenure/120	0.049	0.097*
	(0.037)	(0.043)
R^2	0.253	0.277
No. of observations	247,140	247,140

Notes: the experience measure is actural work experience in months. All specifications control for years of education, year effects, urban residence, and region of residence. Specifications with experience also control for a quadratic term in actural experience and education interacted with experience, and specifications with tenure control for a quadratic term in tenure and education interacted with tenure. The numbers in parentheses are White/Huber standard errors that account for multiple observations per person. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

for non-college graduates.

$$\begin{aligned}\ln w_i = \beta_0 &+ \beta_S S_i + \beta_{S,X}(S_i \times X_i) + \beta_{S,T}(S_i \times T_i) \\ &+ \beta_{AFQT} AFQT_i + \beta_{AFQT,X}(AFQT_i \times X_i) + \beta_{AFQT,T}(AFQT_i \times T_i) \\ &+ \beta_{Black} Black_i + \beta_{Black,X}(Black_i \times X_i) + \beta_{Black,T}(Black_i \times T_i) \\ &+ \beta_\Omega \Omega_i + H(X_i) + \epsilon_i, \end{aligned} \tag{3}$$

The main coefficients of interest are $\beta_{AFQT,X}$, the coefficient on AFQT-experience interaction term, and $\beta_{Black,X}$, the coefficient on *Black*-experience interaction term. Purely asymmetric learning predicts that $\beta_{AFQT,X}$ and $\beta_{Black,X}$ should be equal to zero, and non-purely asymmetric learning indicates non-zero coefficients on experience interaction terms.

We report the estimates of Equation (3) in column (2) of Table 5 for non-college graduates considering that our previous results indicate limited scope for employer learning in the market for college graduates. For ease of comparison, we present in column (1) the results from column (2) of Table 3. When both tenure and experience interactions are included as regressors, the coefficient on the AFQT-experience interaction term is not statistically different from zero, whereas the coefficient on the AFQT-tenure interaction remains significant, providing empirical evidence in favor of purely asymmetric learning. This finding suggests that outside firms have limited access to information about workers' productivity as measured by AFQT over time.

When both black-tenure and black-experience interaction terms are included in the wage Equation (3), the initial negative black coefficient becomes smaller (in absolute value) but remains statistically significant. The significantly positive coefficient on the black-tenure interaction indicates that current firm learns about black workers' productivity over time and rely less on the race information to infer their productivity. The significantly negative coefficient on black-experience interaction is consistent with outside firms not learning about black workers' true productivity over time. Black workers without a college degree appear to be more discriminated on jobs requiring more work experience. These results provide supporting evidence for the assumption of purely asymmetric learning for non-college educated workers.

6.3 Occupation and Industry

Workers of different demographic characteristics and skills sort themselves into different sectors in the labor market (Heckman and Sedlacek (1985)). If black and white workers sort themselves into jobs that require different skill levels or sectors that pay different wages, then the observed wage differences may be due to factors different from those implied by the learning-based statistical discrimination model. One alternative explanation is that black workers are more likely to be hired into jobs and sectors that pay lower wages at the start of their career and to be trapped in such jobs. The initial job assignments and sector allocations could influence the menu of workers' career paths. The evidence of statistical discrimination could be due to differential job sorting by black and white workers.[15]

To test the possibility that racial wage gap is driven by blacks and whites being sorted into jobs of different skill levels, we add initial occupation to Equations (1) and (2) as an additional control and repeat the empirical analysis separately for non-college graduates and college graduates.[16] The regression results are presented in Table 6. In the non-college market (columns (1) and (2)), we find evidence of asymmetric employer learning and statistical discrimination even after controlling for the initial occupations of black and white workers. Wages become more correlated

[15]Racial differences in the initial job assignments and sector allocations could also be an outcome of discrimination, which will strengthen our results.

[16]We distinguish seven occupations: professional workers, managers, sales workers, clerical workers, craftsman and operatives, agricultural labors, and service workers.

Table 6: Estimates Controlling for Initial Occupation

	Non-College Graduates		College Graduates	
	(1)	(2)	(3)	(4)
AFQT	0.034**	0.033**	0.091***	0.141***
	(0.013)	(0.012)	(0.027)	(0.025)
AFQT×experience/120	0.038***		0.040	
	(0.011)		(0.024)	
AFQT×tenure/120		0.075***		-0.076
		(0.020)		(0.041)
Black	-0.066**	-0.140***	0.133**	0.105*
	(0.025)	(0.023)	(0.048)	(0.047)
Black×experience/120	-0.033		-0.109**	
	(0.021)		(0.040)	
Black×tenure/120		0.067		-0.221**
		(0.041)		(0.075)
Initial occupation	Yes	Yes	Yes	Yes
R^2	0.276	0.271	0.318	0.310
No. of observations	189,120	189,120	65,376	65,376

Notes: the experience measure is actural work experience in months. All specifications control for years of education, year effects, urban residence, and region of residence. Specifications with experience also control for a quadratic term in actural experience and education interacted with experience, and specifications with tenure control for a quadratic term in tenure and education interacted with tenure. The numbers in parentheses are White/Huber standard errors that account for multiple observations per person. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

with AFQT over time, and employer learning is faster over job tenure than over actual experience. The *Black* coefficient is initially negative and significant and rises (but insignificantly) with tenure, providing evidence that our main results cannot be attributed to differences in occupation sorting of different racial groups. Including the initial occupation in the regressions also does not alter the results for college graduates presented in columns (3) and (4). College-educated blacks earn an initial wage premium conditional on their AFQT and initial occupation.[17]

We also explore the role of sector allocation by examining the effect of initial industry on the observed racial wage gap.[18] We repeat the empirical analysis for

[17]The results shown in Table 6 provide evidence for race-based statistical discrimination within occupations. Mansour (2012) finds a substantial variation in the time path of black coefficients across occupations. Therefore, the extent of racial statistical discrimination may vary across occupations.

[18]We distinguish twelve industries: agriculture; mining; construction; manufacturing; transportation, communication, and utilities; wholesale and retail trade; finance, insurance, and real estate; business and repair services; personnel services; entertainment and recreation services; professional

Table 7: Estimates Controlling for Initial Industry

	Non-College Graduates		College Graduates	
	(1)	(2)	(3)	(4)
AFQT	0.042**	0.042***	0.108***	0.150***
	(0.013)	(0.012)	(0.026)	(0.025)
AFQT×experience/120	0.036***		0.042	
	(0.011)		(0.024)	
AFQT×tenure/120		0.068***		-0.052
		(0.020)		(0.043)
Black	-0.072**	-0.146***	0.124**	0.101*
	(0.024)	(0.023)	(0.047)	(0.047)
Black×experience/120	-0.036		-0.097*	
	(0.020)		(0.040)	
Black×tenure/120		0.057		-0.201**
		(0.041)		(0.077)
Initial industry	Yes	Yes	Yes	Yes
R^2	0.286	0.281	0.317	0.305
No. of observations	189,120	189,120	65,376	65,376

Notes: the experience measure is actural work experience in months. All specifications control for years of education, year effects, urban residence, and region of residence. Specifications with experience also control for a quadratic term in actural experience and education interacted with experience, and specifications with tenure control for a quadratic term in tenure and education interacted with tenure. The numbers in parentheses are White/Huber standard errors that account for multiple observations per person. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

the two educational groups of interest with initial industry included as a control and present the results in Table 7. The results resemble those without the inclusion of initial industry in Table 3. We also control for initial occupation and industry simultaneously (results not presented but available from the authors), and the main results are not affected. Therefore, the racial wage gap cannot be explained by the variations in workers' initial occupation or industry.[19]

7 Conclusion

In this paper, we combine elements of employer learning and statistical discrimination theories to empirically examine whether employers statistically discriminate

and related services; and public administration.

[19]Tables 6 and 7 present the *OLS* estimates of the wage regressions. Results from the *IV* estimates treating actual experience or job tenure as endogenous are similar and available upon request.

against black workers at time of hiring under different scenarios of employer learning.

Our estimation results show that non-college graduates and college graduates are associated with different patterns of employer learning. At the time of initial hire, employers have to rely on some easily observable characteristics to estimate the productivity of non-college graduates, and they gradually update their expectations as they acquire additional information. The time paths of racial wage gap in the non-college market indicate that employers use race as information to infer workers' productivity and black workers are statistically discriminated. We find that learning correlates more with tenure than with experience. This finding supports the hypothesis that learning is asymmetric in the non-college labor market. Among college graduates, we do not find evidence that black workers are statistically discriminated.

Many statistical discrimination models are built on the assumption that the signal of productivity employers receive from black workers is less reliable than that from white workers at the time of initial hire.[20] Pinkston (2006) applies the framework of employer learning to test this hypothesis, and his estimation results provide evidence supporting this view.[21] An interesting topic for future research is to relax the assumption of equally informative signals from different racial groups and investigate its effect on employer learning and statistical discrimination.

Finally, our paper, as all of the related literature, is not designed to measure to what extent the persistent racial wage differences are due to statistical versus "taste-based" discrimination (in the sense of Becker (1971)).[22] Disentangling the different sources of group inequality remains an important topic for future research.

References

Aigner, Dennis J and Glen G Cain, "Statistical theories of discrimination in labor markets," *ILR Review*, 1977, *30* (2), 175–187. (Cited on pages 2 and 20)

Akerlof, George, "The market for lemons: qualitative uncertainty and the market mechanism," *Quarterly Journal of Economics*, 1970, *84* (3). (Cited on page 1)

Altonji, Joseph G. and Charles R. Pierret, "Employer Learning and Statistical Discrimination," *The Quarterly Journal of Economics*, 2001, *116* (1), 313–350. (Cited on pages 1 and 5)

[20]See e.g., Aigner and Cain (1977), Lundberg and Startz (1983), Cornell and Welch (1996), and Lang (1986).

[21]Flabbi et al. (2016) provide empirical support to the hypothesis that signal quality differs by gender.

[22]See Neal and Johnson (1996) and surveys by Altonji and Blank (1999) and Lang and Lehmann (2012), among others.

Altonji, Joseph G and Rebecca M Blank, "Race and gender in the labor market," *Handbook of labor economics*, 1999, *3*, 3143–3259. (Cited on page 20)

Altonji, Joseph G. and Robert A. Shakotko, "Do Wages Rise with Job Seniority?," *The Review of Economic Studies*, 1987, *54* (3), 437–459. (Cited on page 14)

Arcidiacono, Peter, "Affirmative action in higher education: How do admission and financial aid rules affect future earnings?," *Econometrica*, 2005, *73* (5), 1477–1524. (Cited on page 13)

_, Patrick Bayer, and Aurel Hizmo, "Beyond signaling and human capital: Education and the revelation of ability," *American Economic Journal: Applied Economics*, 2010, *2* (4), 76–104. (Cited on pages 2, 3, 7, 10, 11, and 13)

Arrow, Kenneth, "The Theory of Discrimination," in Orley Ashenfelter and Albert Rees, eds., *Discrimination in Labor Markets*, Princeton, N.J.: Princeton University Press, 1973, pp. 3–33. (Cited on page 2)

Bates, Michael, "Public and Private Employer Learning: Evidence from the Adoption of Teacher Value Added," *Journal of Labor Economics*, 2020, *38* (2), 375–420. (Cited on page 4)

Bauer, Thomas K and John P Haisken-DeNew, "Employer learning and the returns to schooling," *Labour Economics*, 2001, *8* (2), 161–180. (Cited on page 4)

Becker, Gary, *The economics of discrimination*, Chicago: The University of Chicago Press (first ed. 1957), 1971. (Cited on page 20)

Bernhardt, Dan, "Strategic promotion and compensation," *The Review of Economic Studies*, 1995, *62* (2), 315–339. (Cited on page 2)

Coate, Stephen and Glenn C. Loury, "Will Affirmative-Action Policies Eliminate Negative Stereotypes?," *The American Economic Review*, December 1993, *83* (5), 1220–1240. (Cited on page 2)

Cornell, Bradford and Ivo Welch, "Culture, information, and screening discrimination," *Journal of Political Economy*, 1996, pp. 542–571. (Cited on page 20)

Eaton, Morris L., *Multivariate Statistics : A Vector Space Approach* Wiley series in probability and mathematical statistics, New York : Wiley, 1983. (Cited on page 2)

Fan, Xiaodong and Jed DeVaro, “Job Hopping and Adverse Selection in the Labor Market,” *The Journal of Law, Economics, and Organization*, 12 2020, *36* (1), 84–138. (Cited on page 4)

Fang, Hanming and Andrea Moro, “Theories of Statistical Discrimination and Affirmative Action: A Survey,” in Jess Benhabib, Matthew O. Jackson, and Alberto Bisin, eds., *Handbook of Social Economics*, Vol. 1A, The Netherlands: North Holland, 2011, chapter V, pp. 133–200. (Cited on page 2)

Farber, Henry S. and Robert Gibbons, “Learning and Wage Dynamics,” *The Quarterly Journal of Economics*, 1996, *111* (4), 1007–1047. (Cited on page 1)

Flabbi, Luca, Mario Macis, Andrea Moro, and Fabiano Schivardi, “Do Female Executives Make a Difference? The Impact of Female Leadership on Gender Gaps and Firm Performance,” Working Paper 22877, National Bureau of Economic Research December 2016. (Cited on page 20)

Gibbons, Robert and Lawrence F Katz, “Layoffs and lemons,” *Journal of labor Economics*, 1991, *9* (4), 351–380. (Cited on page 4)

Golan, Limor, “Counteroffers and efficiency in labor markets with asymmetric information,” *Journal of Labor Economics*, 2005, *23* (2), 373–393. (Cited on page 2)

_ , “Wage signaling: A dynamic model of intrafirm bargaining and asymmetric learning,” *International Economic Review*, 2009, *50* (3), 831–854. (Cited on page 2)

Greenwald, Bruce C., “Adverse selection in the labour market,” *The Review of Economic Studies*, 1986, *53* (3), 325–347. (Cited on pages 2 and 1)

Heckman, James J and Guilherme Sedlacek, “Heterogeneity, aggregation, and market wage functions: an empirical model of self-selection in the labor market,” *Journal of Political Economy*, 1985, *93* (6), 1077–1125. (Cited on page 17)

Kahn, Lisa B, “Asymmetric Information Between Employers,” *American Economic Journal: Applied Economics*, 2013, *5* (4), 165–205. (Cited on page 4)

Lang, Kevin, “A language theory of discrimination,” *The Quarterly Journal of Economics*, 1986, pp. 363–382. (Cited on page 20)

_ **and Jee-Yeon K Lehmann**, “Racial discrimination in the labor market: Theory and empirics,” *Journal of Economic Literature*, 2012, *50* (4), 959–1006. (Cited on pages 2 and 20)

— **and Michael Manove**, "Education and labor market discrimination," *American Economic Review*, 2011, *101* (4), 1467–96. (Cited on page 13)

Lange, Fabian, "The Speed of Employer Learning," *Journal of Labor Economics*, 2007, *25* (1), 1–35. (Cited on pages 2 and 5)

Lesner, Rune V, "Testing for Statistical Discrimination based on Gender," *Labour*, 2018, *32* (2), 141–181. (Cited on page 2)

Lundberg, Shelly J and Richard Startz, "Private discrimination and social intervention in competitive labor market," *The American Economic Review*, 1983, *73* (3), 340–347. (Cited on page 20)

Mansour, Hani, "Does employer learning vary by occupation?," *Journal of Labor Economics*, 2012, *30* (2), 415–444. (Cited on pages 2 and 12)

Neal, Derek A and William R Johnson, "The role of premarket factors in black-white wage differences," *Journal of political Economy*, 1996, *104* (5), 869–895. (Cited on pages 13 and 20)

Phelps, Edmund S, "The Statistical Theory of Racism and Sexism," *The American Economic Review*, 1972, *62* (4), 659–661. (Cited on pages 2 and 1)

Pinkston, Joshua C., "A Test of Screening Discrimination with Employer Learning," *Industrial and Labor Relations Review*, 2006, *59* (2), 267–284. (Cited on page 20)

— , "A Model of Asymmetric Employer Learning with Testable Implications," *The Review of Economic Studies*, 2009, *76* (1), 367–394. (Cited on pages 2, 4, and 5)

Schönberg, Uta, "Testing for Asymmetric Employer Learning," *Journal of Labor Economics*, 2007, *25* (4), 651–691. (Cited on pages 3 and 15)

Waldman, Michael, "Job assignments, signalling, and efficiency," *The RAND Journal of Economics*, 1984, *15* (2), 255–267. (Cited on page 2)

External Appendix to: Testing for Asymmetric Employer Learning and Statistical Discrimination

By Suqin Ge, Andrea Moro, and Beibei Zhu

Extending the Farber and Gibbons (1996)-Altonji and Pierret (2001) model to the asymmetric employer learning case to study statistical discrimination is not straightforward for two reasons. First, Altonji and Pierret (2001)'s identification strategy relies on employers updating continuously workers' wages based on new information arising over time. But, if information is private to the employer, one has to model why wages change. Solutions have been proposed in the literature cited in the introduction, but no comprehensive theory has emerged so far. The second reason is that one also has to model how and why job changes occur. In a "competitive" environment, workers that are willing to move have average productivity below the offer, which makes the wage offer unprofitable to the new employer. This may collapse the market for job changes by a process similar to Akerlof (1970)'s market for lemons.[23]

Our goal in this appendix is not to present a model of the labor market under asymmetric information that provides a general solution to these issues, but to highlight forces that we believe play a crucial role in determining wages under incomplete information and derive empirical implications that plausibly hold under less stringent assumptions. Therefore, we derive empirical implications from a standard signal extraction model following two admittedly strong assumptions: (i) wages are not equal to expected marginal product, but are equal to a fraction of the expected surplus that does not change with tenure.[24] For example, wages could result from the Generalized Nash-bargaining solution of a cooperative game with employers and workers bargaining over a share of the (employers-expected) surplus. (ii) we assume for simplicity that separations occur exogenously for labor-market related reasons that are beyond the influence of their employers.

We start our analysis extending the standard statistical discrimination model established by Phelps (1972) to include a time dimension to allow for employers' learning. Firms compete for workers and maximize output given wages. Workers care only about wages. Workers draw their productivity q from a normal distribution with mean $\mu(X)$ and variance $\sigma^2(X)$, where X is a set of variables observed by the

[23]To support the equilibrium, for example, Greenwald (1986) assumes that some, but not all, agents separate exogenously.

[24]Note that even the symmetric learning model relies on final output being not contractible, which is the case for example when workers' contribution to total output cannot be observed with certainty.

employer that correlate with productivity. In the standard statistical discrimination model, X includes group identity such as race or gender. Ability and wages are expressed in logarithms to guarantee that they have positive values in levels.

Employers initially observe a signal of productivity s_0 from a new worker, where $s_0 = q + \epsilon_0$. After hiring, in each period t, they observe from the employee an additional signal $s_t = q + \epsilon_t$. All ϵ_t's are independently and normally distributed with mean 0 and variance σ_ϵ^2. The signal's variance can be interpreted as a measure of the signal's information quality (higher variance corresponds to poorer quality).

The analysis is started by exploring the effect of the incumbent employer's learning. In each period, the employer computes the worker's expected productivity given the signals observed. New workers are offered expected productivity $E(q|s_0)$. The standard properties of the bivariate normal distribution[25] imply that:

$$E(q|s_0) = (1-\alpha)\mu(X) + \alpha s_0,$$
$$\alpha = \frac{\sigma^2}{\sigma_\epsilon^2 + \sigma^2}.$$

In this expression, expected productivity is a weighted average of the population average skill and the initial signal, with weights equal to the relative variance of the two variables. When the signal is perfectly informative ($\sigma_\epsilon = 0$), the population mean is ignored; when the signal is pure noise ($\sigma_\epsilon = \infty$), expected ability is equal to the population's average, conditional on observables X. With a partially informative signal, the conditional expected productivity is increasing in q and s_0. The conditional distribution, which we denote with $\phi(q|s_0)$ is also normal, with mean equal to $E(q|s_0)$ and variance $\alpha\sigma_\epsilon^2$.

As the worker increases his tenure with a firm, the incumbent employer exploits information from multiple signals, providing more precise information about true productivity. In period 1, the firm has another signal available, s_1. The normality of the conditional distribution is preserved, and its moments can be derived using the same formula:

$$E(q|s_0, s_1) = (1-k_1)E(q|s_0) + k_1 s_1, \tag{4}$$
$$\text{with } k_1 = \frac{\alpha\sigma_\epsilon^2}{\alpha\sigma_\epsilon^2 + \sigma_\epsilon^2} = \frac{\alpha}{\alpha+1}$$

k_T is used to denote the weight assigned to the last signal s_T in the computation of the expected productivity, and α_T is the weight assigned to the average value

[25] See Eaton (1983).

of all signals received up to period T, which we denote with $\overline{s}$. We can write the expected productivity in Expression (4) as a function of the average value of all signals received so far:

$$\begin{aligned}
E(q|s_0,s_1) &= (1-k_1)\left((1-\alpha)\mu+\alpha s_0\right)+k_1 s_1\\
&= \frac{1}{\alpha+1}(1-\alpha)\mu+\frac{1}{\alpha+1}\alpha(s_0+s_1)\\
&= \frac{1-\alpha}{1+\alpha}\mu+\frac{2\alpha}{1+\alpha}\frac{\sum s_i}{2}=\\
&= \left(1-\frac{2\alpha}{1+\alpha}\right)\mu+\frac{2\alpha}{1+\alpha}\frac{\sum s_i}{2}\\
\left(\text{define } \alpha_1\equiv\frac{2\alpha}{1+\alpha}\Rightarrow\right) &= (1-\alpha_1)\mu+\alpha_1\overline{s} \text{ with}
\end{aligned}$$

In period 2, another signal becomes available, s_2, we follow the same steps:

$$\begin{aligned}
E(q|s_0,s_1,s_2) &= (1-k_2)E(q|s_0,s_1)+k_2 s_2,\\
\text{with } k_2 &= \frac{\frac{\alpha\sigma_\epsilon^2}{1+\alpha}}{\frac{\alpha\sigma_\epsilon^2}{1+\alpha}+\sigma_\epsilon^2}=\frac{\alpha}{2\alpha+1}\\
&= (1-k_2)(1-\alpha_1)\mu+(1-k_2)\alpha_1\frac{s_1+s_2}{2}+k_2 s_2\\
&= \frac{1}{2\alpha+1}(1-\alpha)\mu+\frac{3\alpha}{2\alpha+1}\frac{s_1+s_2+s_3}{3}\\
\left(\text{define } \alpha_2\equiv\frac{3\alpha}{1+2\alpha}\Rightarrow\right) &= (1-\alpha_2)\mu+\alpha_2\overline{s}
\end{aligned}$$

Hence, by induction, after T periods and $T+1$ signals,

$$\begin{aligned}
E(q|s_0,s_1\dots s_T) &= (1-\alpha_T)\mu+\alpha_T\overline{s} \text{ with } \alpha_T=\frac{(T+1)\alpha}{1+T\alpha} \qquad (5)\\
Var(q|s_0,s_1...s_T) &= \frac{\alpha\sigma_\epsilon^2}{1+T\alpha}
\end{aligned}$$

For the purpose of comparing workers with continuous and discontinuous tenure histories, note that from (5), we can calculate the weight the employer places on the unconditional mean μ :

$$\begin{aligned}
1-\alpha_T &= 1-\frac{(T+1)\alpha}{1+T\alpha}=\frac{1-\alpha}{1+T\alpha}\\
&= \frac{\sigma_\epsilon^2}{(T+1)\sigma^2+\sigma_\epsilon^2}
\end{aligned}$$

Note that the weight α_T placed on the signals average is increasing in T and converges to 1. As tenure increases, the worker's expected productivity gets closer to her true productivity.

As described at the beginning of this Appendix, we assume wages are a fraction of expected worker's productivity determined by a Generalized Nash-Bargaining Solution of a cooperative game between employers and worker in which the bargaining power parameter does not change with time. Hence, any correlation between observables and wages can be determined by looking at their correlation with expected productivity.

Consider an econometrician observing wages and a one-time signal of skill r not observed by the employer, such that $r = q + \epsilon_r,\ \epsilon_r \sim N(0, \sigma_r^2)$. In the empirical section, we assume the AFQT to be one such signal, therefore we will refer to signal r as AFQT in the following discussion. We can compute the covariance of this signal with expected productivity:

$$
\begin{aligned}
Cov(r, E(q|s_0 ... s_T)) &= Cov\left(q + \epsilon_r, (1 - \alpha_T)\mu(X) + \alpha_T \left(\frac{\sum_{t=0}^{T}(q + \epsilon_t)}{T+1}\right)\right) \\
&= \alpha_T Var(q),
\end{aligned}
$$

that is, as the number of signals increases, the expected productivity increasingly covaries with the signal observed by the econometrician.

Empirical Implication 1. Under the assumptions of the model, in a wage regression, the interaction of workers' tenure with AFQT displays a positive coefficient.

The result does not rely necessarily on assuming perfect competition in the labor market. For example, if employers and workers bargain over a share of the (expected) surplus, a sufficient condition for the implication to hold is that the bargaining power of the worker does not change considerably with tenure.[26]

Consider now workers hired by a new employer immediately before the beginning of period T. We assume for simplicity that separations occur exogenously for labor-market related reasons that are beyond the influence of their employers.

The new employer collects information from the worker's resume and other signals, which we summarize with a new signal observed by the new employer in period T labeled $\nu = q + \epsilon_v$, with $\epsilon_\nu \sim N(0, \sigma_\nu^2)$. The information available to the new employer is at least as good as the first signal received by the first employer, such

[26]Note that the model relies on final output being not contractible, which is the case for example when workers' contribution to total output cannot be observed with certainty.

that $\sigma_\nu \leq \sigma_\epsilon$. The expected productivity given this information is also normal:

$$E(q|\nu) = (1-\alpha^N)\mu(X) + \alpha^N \nu \tag{6}$$
$$\text{with } \alpha^N = \frac{\sigma^2}{\sigma_\nu^2 + \sigma^2}$$

Crucially, we allow that the information available to the new employer at the time of hiring is worse than the information available to the current employer:

$$\alpha^N < \alpha_{T-1} \Leftrightarrow \sigma_\nu^2 > \frac{\sigma_\epsilon^2}{T} \tag{7}$$

The expressions are equivalent because they assume that the variance of the signal for the new employer is greater than the variance of the average signal of the incumbent. We believe this condition to be realistic because worker's resume, job interviews, and aptitude tests cannot substitute from day-to-day interactions over the workers' tenure.[27] If the new employer does not infer any information from prior job history, then her signal ν carries the same information as any other signal available to new employers, or $\sigma_\nu = \sigma_\epsilon$, implying $\alpha^N = \alpha$, which is less than α_{T-1}, a situation we label as *purely asymmetric learning.*

After hiring, the new employer's expectations as the previous employer's: the new employer extracts a signal every period from the same distribution as the previous employer's signals and revises her posterior expectations using Bayes' rule.

For example, two workers, Mary and John, have the same experience T. Mary has always been with the same employer, whereas John has worked for two employers, changing job after $T-1$ periods, and has stayed for one period with the new employer. We can compare the expected productivities for the two workers. We use superscript

[27] In a model where outside employers bid with current employers for workers' wages in an auction, Pinkston (2009) proves that the incumbent employer's information about workers is completely revealed to the outside employer after the bidding process. For our purposes, we assume that frictions exist in the environment that prevent the information to be completely revealed to competing employers, or that unemployment spells prevent the bidding process to completely reveal the information. Potential employers could also observe wages and job history, therefore learning some of the information available to current employers. However, there are always other forms of compensation besides wages whose value to the worker is hard to assess for an outside employer.

N to denote parameters related to the worker moving to a new employer:

$$\text{Mary:} \qquad E(q|s_0,\dots,s_T) = (1-\alpha_T)\mu + \alpha_T \sum_{t=0}^{T} \frac{s_t}{T+1}$$
$$\text{with } \alpha_T = \frac{(T+1)\alpha}{1+T\alpha}$$
$$\text{John:} \qquad E(q|\nu,s_T) = (1-\beta^N)\left((1-\alpha^N)\mu + \alpha^N\nu\right) + \beta^N s_T \tag{8}$$
$$\text{with } \beta^N = \frac{\alpha^N\sigma_v^2}{\alpha^N\sigma_v^2+\sigma_\epsilon^2}.$$

Note that the weight placed on the conditional mean $\mu(X)$ in evaluating the expected productivity is $(1-\alpha_T)$ for Mary and $(1-\beta^N)(1-\alpha^N)$ for John.

One period after the new hire the new employers observes signal $s_T = q+\epsilon$, with:

$$E(q|\nu,s_T) = (1-\beta^N)\left((1-\alpha^N)\mu + \alpha^N\nu\right) + \beta^N s_T$$
$$\beta^N = \frac{\alpha^N\sigma_\nu^2}{\alpha^N\sigma_\nu^2+\sigma_\epsilon^2}$$

We compute the weight the new employer places on the unconditional average μ :

$$\begin{aligned}
(1-\beta^N)(1-\alpha^N) &= \left(1-\frac{\alpha^N\sigma_\nu^2}{\alpha^N\sigma_\nu^2+\sigma_\epsilon^2}\right)\left(1-\frac{\sigma^2}{\sigma^2+\sigma_\nu^2}\right) \\
&= \frac{\sigma_\epsilon^2}{\alpha^N\sigma_\nu^2+\sigma_\epsilon^2}\frac{\sigma_\nu^2}{\sigma^2+\sigma_\nu^2} = \frac{\sigma_\epsilon^2}{\frac{\sigma^2}{\sigma^2+\sigma_\nu^2}\sigma_\nu^2+\sigma_\epsilon^2}\frac{\sigma_\nu^2}{\sigma^2+\sigma_\nu^2} = \\
&= \frac{\left(\sigma^2+\sigma_\nu^2\right)\sigma_\epsilon^2}{\sigma^2\sigma_\nu^2+\left(\sigma^2+\sigma_\nu^2\right)\sigma_\epsilon^2}\frac{\sigma_\nu^2}{\sigma^2+\sigma_\nu^2} = \frac{\sigma_\nu^2\sigma_\epsilon^2}{\sigma^2\sigma_\nu^2+\left(\sigma^2+\sigma_\nu^2\right)\sigma_\epsilon^2} \\
&= \frac{\sigma_\epsilon^2}{\sigma^2+\frac{\sigma^2\sigma_\epsilon^2}{\sigma_\nu^2}+\sigma_\epsilon^2} > \frac{\sigma_\epsilon^2}{\sigma^2+\frac{\sigma^2\sigma_\epsilon^2}{\sigma_\epsilon^2/T}+\sigma_\epsilon^2} = \frac{\sigma_\epsilon^2}{(T+1)\sigma^2+\sigma_\epsilon^2} = (1-\alpha_T),
\end{aligned}$$

where the inequality follows from substituting σ_ν^2 with (by assumption in (7)) a smaller number σ_ϵ^2/T. Therefore, incumbent with T signals places a smaller weight on the unconditional mean μ than the employer of a worker that switched after $T-1$ periods. A similar derivation follows for subsequent Q periods:

$$(1-\beta_Q^N)(1-\alpha^N) > (1-\alpha_{T+Q})$$

$$\beta_Q^N = \frac{(Q+1)\beta^N}{1+Q\beta^N},$$

that is, additional signals improve the quality of information for the new employer; however, the quality of information never catches up with the information of the employer of a worker with same experience but uninterrupted tenure.

Assuming again that the econometrician has a signal of productivity such as the AFQT $r = q + \epsilon_r$, we can compute

$$\begin{aligned}
\text{Mary:}\quad Cov\left(r, E(q|s_0, ...s_T)\right) =& Cov\left(q + \epsilon_r, (1-\alpha_T)\mu + \alpha_T \sum_{t=0}^{T} \frac{q+\epsilon_t}{T+1}\right) \\
=& \alpha_T Var(q)
\end{aligned} \tag{9}$$

$$\begin{aligned}
\text{John:}\quad Cov\left(r, E(q|\nu, s_T)\right) =& Cov(q + \epsilon_r, (1-\beta^N)\left((1-\alpha^N)\mu + \alpha^N \nu\right) + \beta^N(q + \epsilon_T)) \\
=& \left((1-\beta^N)\alpha^N + \beta^N\right) Var(q)
\end{aligned} \tag{10}$$

We have shown $(1-\beta^N)(1-\alpha^N) > 1-\alpha_T$, therefore $\alpha_T > (1-\beta^N)\alpha^N + \beta^N$. Hence, the wages of workers with discontinuous work histories covary with the econometricians' signals of productivity less than workers with continuous work histories. But because tenure is always less than experience, we can conclude the following:

Empirical Implication 2. Consider two regressions that include the interactions of either tenure with AFQT or experience with AFQT, if learning is asymmetric, then the coefficient of the interaction of tenure with AFQT is positive and larger than the coefficient of the interaction of experience with AFQT.

Some workers with high experience have low tenure, therefore the correlation of experience with AFQT is lower. Note that the opposite implication would be true if learning is symmetric. In that case, some workers with low tenure have high experience. Their employers had the opportunity to learn more, therefore the coefficient on tenure should be attenuated relative to the coefficient on experience.

We extend the model to study its implications on statistical discrimination. The two groups of workers with easily recognizable traits are minority (M) and dominant (D) groups. Assume that $\mu(M) < \mu(D)$ and that employers use race for labor market decisions,[28] the productivity signals observed by the econometrician and the employer, such as racial identity, are accounted for by the term $\mu(X)$; therefore variables in X will be less correlated with wages over time.

Empirical Implication 3. Under the assumptions of the model, in a wage regression including a race M dummy, if group M is statistically discriminated against, then

[28] We focus here on the empirical implications of such behavior ignoring its legal aspects: using race even for informational purposes is in general illegal, but employers may be able to do so by using other proxies for race. Ultimately, whether or not employers statistically discriminate is an empirical question.

the coefficient on such dummy is negative, but its interaction with tenure is positive so that the negative effect declines over time.

This implication would hold also using experience instead of tenure because experience and tenure are positively correlated. However, in the extreme case of purely asymmetric learning, the interaction of the race dummy with tenure should display stronger effects (see Table 3, columns 1 and 2, and Table 5, column 2).